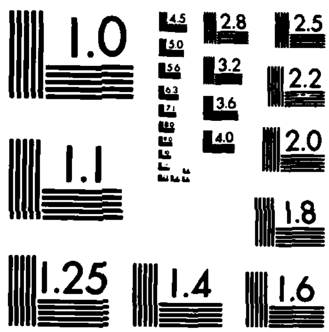
AD-A120 968 DROPPING OBSERVATIONS WITHOUT AFFECTING POSTERIOR AND PREDICTIVE DISTRIBU..(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER N R DRAPER ET AL. AUG 82 1/1

UNCLASSIFIED MRC-TSR-2409 DAAG29-80-C-0041 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A120968

MRC Technical Summary Report #2409

DROPPING OBSERVATIONS WITHOUT AFFECTING POSTERIOR AND PREDICTIVE DISTRIBUTIONS

Norman R. Draper and Irwin Guttman

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

August 1982

(Received June 1, 1982)

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

DROPPING OBSERVATIONS WITHOUT AFFECTING POSTERIOR AND PREDICTIVE DISTRIBUTIONS

Norman Draper and Irwin Guttman*

Technical Summary Report #2409

August 1982

ABSTRACT

Suppose in a distribution problem, the sample information $\underset{\sim}{W}$ is split into two pieces $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$, and the parameters involved are split into two sets, $\underset{\sim}{\phi}$ containing the parameters of interest, and $\underset{\sim}{\theta}$ containing nuisance parameters. It is shown that, under certain conditions, the posterior distribution of $\underset{\sim}{\phi}$ does not depend on the data $\underset{\sim}{W}_2$, which can thus be ignored. This also has consequences for the predictive distribution of future (or missing) observations. In fact, under similar conditions, the predictive distributions using $\underset{\sim}{W}$ or just $\underset{\sim}{W}_1$ are identical.

AMS (MOS) Subject Classifications: 62F15

Key Words: Bayes' theorem, Nuisance parameters, Predictive distribution

Work Unit Number 4 (Statistics and Probability)

*University of Toronto, Ontario, Canada.

Sponsored by the United States Army under Contract Nos. DAAG29-80-C-0041, and DAAG29-80-C-0113.

SIGNIFICANCE AND EXPLANATION

In the application of Bayesian methods, some posterior and predictive distributions may be unaffected when portions of the data are ignored. Conditions under which this is true are given, and examples are provided.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not the authors of this report.

DROPPING OBSERVATIONS WITHOUT AFFECTING POSTERIOR AND

PREDICTIVE DISTRIBUTIONS

Norman Draper and Irwin Guttman*

1. MAIN RESULT

Suppose, that $\underset{\sim}{W} = (\underset{\sim}{W}_1', \underset{\sim}{W}_2')'$ is a vector of random variables whose density function $f_{\underset{\sim}{W}}(\underset{\sim}{w}|\underset{\sim}{\phi},\underset{\sim}{\theta})$ depends on two sets of parameters $\underset{\sim}{\phi},\underset{\sim}{\theta}$. Suppose further that we are interested in $\underset{\sim}{\phi}$, that $\underset{\sim}{\theta}$ will be regarded as a vector of nuisance parameters and that the following conditions hold.

1. $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$ are statistically independent.
2. The marginal distribution $f_{\underset{\sim}{W}_2}(\underset{\sim}{w}_2|\underset{\sim}{\theta})$ of W_2 depends only on $\underset{\sim}{\theta}$ and not on $\underset{\sim}{\phi}$.
3. The marginal distribution of $\underset{\sim}{W}_1 = (\underset{\sim}{W}_{11}', \underset{\sim}{W}_{12}')'$ is such that

$$f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1|\underset{\sim}{\phi},\underset{\sim}{\theta}) = f_{\underset{\sim}{W}_{11}}(\underset{\sim}{w}_{11}|\underset{\sim}{\theta})\, f_{\underset{\sim}{W}_{12}|\underset{\sim}{W}_{11}}(\underset{\sim}{w}_{12}|\underset{\sim}{w}_{11};\underset{\sim}{\phi}) \tag{1.1}$$

4. The prior information about the parameter sets $\underset{\sim}{\phi}$ and $\underset{\sim}{\theta}$ is such that

$$p(\underset{\sim}{\phi},\underset{\sim}{\theta}) = a(\underset{\sim}{\theta})b(\underset{\sim}{\phi}) \tag{1.2}$$

so that $\underset{\sim}{\theta}$, $\underset{\sim}{\phi}$ are independent <u>a priori</u>.

<u>Theorem 1</u>. Under conditions 1-4, the <u>marginal posterior of</u> $\underset{\sim}{\phi}$ <u>based on</u> $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$ <u>does not depend on</u> $\underset{\sim}{W}_2$.

<u>Proof</u>. The posterior of $\underset{\sim}{\phi}$ given $(\underset{\sim}{W}_1,\underset{\sim}{W}_2)' = (\underset{\sim}{w}_1,\underset{\sim}{w}_2)'$ is

$$p(\underset{\sim}{\phi}|\underset{\sim}{w}_1,\underset{\sim}{w}_2) \propto \int_{\underset{\sim}{\theta}} p(\underset{\sim}{\phi},\underset{\sim}{\theta}) f_{\underset{\sim}{W}_1,\underset{\sim}{W}_2}(\underset{\sim}{w}_1,\underset{\sim}{w}_2|\underset{\sim}{\phi},\underset{\sim}{\theta})\,d\underset{\sim}{\theta} \tag{1.3a}$$

$$\propto b(\underset{\sim}{\phi}) \int_{\underset{\sim}{\theta}} a(\underset{\sim}{\theta}) f_{\underset{\sim}{W}_2}(\underset{\sim}{w}_2|\underset{\sim}{\theta}) f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1|\underset{\sim}{\phi},\underset{\sim}{\theta})\,d\underset{\sim}{\theta} \tag{1.3b}$$

*University of Toronto, Ontario, Canada.

Sponsored by the United States Army under Contract Nos. DAAG29-80-C-0041 and DAAG29-80-C-0113.

$$\propto b(\underset{\sim}{\phi}) f_{\underset{\sim}{W}_{12}|\underset{\sim}{W}_{11}}(\underset{\sim}{w}_{12}|\underset{\sim}{w}_{11};\underset{\sim}{\phi}) \times$$

$$\int_{\underset{\sim}{\theta}} a(\underset{\sim}{\theta}) f_{\underset{\sim}{W}_2}(\underset{\sim}{w}_2|\underset{\sim}{\theta}) f_{\underset{\sim}{W}_{11}}(\underset{\sim}{w}_{11}|\underset{\sim}{\theta}) d\underset{\sim}{\theta}. \quad (1.3c)$$

The integral in (1.3c) is clearly a function of $\underset{\sim}{w}_2$ and $\underset{\sim}{w}_{11}$, i.e., constant with respect to $\underset{\sim}{\phi}$, and so may be absorbed into the constant of proportionality. This proves the result stated.

2. TWO EXAMPLES

Example 2.1. Consider the bivariate normal distribution with vector of means $\underset{\sim}{\mu} = (\mu_1,\mu_2)'$, variance-covariance matrix $\underset{\sim}{\Sigma} = ((\sigma_{ij}))$, and inverse $\underset{\sim}{\Sigma}^{-1} = ((c_{ij}))$. Let

$$\begin{array}{lcl}
\alpha = \mu_2 + c_{12}\mu_1/c_{22}, & & \mu_1 = \eta_1, \\
\beta = -c_{21}/c_{22}, & & \mu_2 = \alpha + \beta\eta_1, \\
d_{11} = c_{22}, & \text{or} & c_{11} = d_{22} + \beta^2 d_{11}, \quad (2.1) \\
d_{22} = c_{11} - c_{21}^2/c_{22}, & & c_{22} = d_{11}, \\
\eta_1 = \mu_1, & & c_{21} = -\beta d_{11}.
\end{array}$$

We remark in passing that, if (2.1) is considered as a transformation from $(\mu_1,\mu_2,c_{11},c_{22},c_{21})$ to $(\eta_1,\alpha,\beta,d_{11},d_{22})$, then the Jacobian has absolute value d_{11}.

Using (2.1), we can re-write the usual form of the bivariate normal frequency function in x_1 and x_2 as

$$f(x_1,x_2|\mu_1,\alpha,\beta,d_{11},d_{22})$$

$$= f(x_2|x_1;\alpha,\beta,d_{11}) \times f(x_1|\eta_1,d_{22}) \quad (2.2)$$

where

$$f(x_2|x_1;\alpha,\beta,d_{11}) = (d_{11}/2\pi)^{1/2} \exp\{-\tfrac{1}{2}d_{11}(x_2-\alpha-\beta x_1)^2\}, \qquad (2.2a)$$

$$f(x_1|\eta_1;d_{22}) = (d_{22}/2\pi)^{1/2} \exp\{-\tfrac{1}{2}d_{22}(x_1-\mu_1)^2\}. \qquad (2.2b)$$

Now suppose that our data on x_1 and x_2 are divided up into two independent portions, satisfying condition1 of Theorem 1:

$\underset{\sim}{W}_1$, consisting of n independent observations (x_{1i},x_{2i}), $i = 1,\ldots,n$, on both x_1 and x_2.

$\underset{\sim}{W}_2$, consisting of n* independent observations x^*_{1j}, $j = 1,\ldots,n^*$, on x_1 alone; the $x^*_{2j} = y_j$ are not observed.

This is the well known missing observations problem (see Draper and Guttman, 1977, and prior references listed therein). In the notation of Section 1, we write

$$\underset{\sim}{\phi} = (\alpha,\beta,d_{11})' \qquad (2.4)$$

for the vector of parameters of interest, and

$$\underset{\sim}{\theta} = (\eta_1,d_{22})' \qquad (2.5)$$

for the vector of nuisance parameters. To obtain a prior for $(\underset{\sim}{\phi},\underset{\sim}{\theta})$, we transform the usual non-informative prior

$$p(\mu_1,\mu_2,c_{11},c_{22},c_{12}) \propto (c_{11}c_{22}-c_{12}^2)^{-3/2} \tag{2.6}$$

using (2.1), remembering to insert the Jacobian. This provides

$$p[(\alpha,\beta,d_{11}),(\mu_1,d_{22})] \propto d_{11}^{-1/2}d_{22}^{-3/2} \tag{2.7}$$

from which it is clear that condition 4 of Theorem 1 is satisfied. If we define

$$\begin{aligned} \underset{\sim}{W}_{11} &= \{x_{1i},\ \ i = 1,2,\ldots,n\}, \\ \underset{\sim}{W}_{12} &= \{x_{2i},\ \ i = 1,2,\ldots,n\}, \end{aligned} \tag{2.8}$$

and employ (2.2) to (2.5), condition 3 of Theorem 1 is satisfied, and condition 2 is satisfied because of (2.2b). It follows that Theorem 1 applies and that the posterior of $\underset{\sim}{\phi}$ given $(\underset{\sim}{W}_1,\underset{\sim}{W}_2)$ depends only on $\underset{\sim}{W}_1$. In fact it can be verified directly that

$$p(\underset{\sim}{\phi}|\underset{\sim}{W}_1,\underset{\sim}{W}_2) \propto d_{11}^{(n-1)/2} \exp[-\tfrac{1}{2}d_{11}\{S+(\underset{\sim}{\phi}_1-\hat{\underset{\sim}{\phi}}_1)\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{\phi}_1-\hat{\underset{\sim}{\phi}}_1)\}], \tag{2.9}$$

where

$$\begin{aligned} S &= \sum_{i=1}^{n} (y_i-\hat{\alpha}-\hat{\beta}x_i)^2, \qquad \underset{\sim}{\phi}_1 = (\alpha,\beta)', \\ \underset{\sim}{X} &= (\underset{\sim}{1},\underset{\sim}{x}_1), \end{aligned} \tag{2.10}$$

$\underset{\sim}{1}$ is an $n \times 1$ vector of ones, $\underset{\sim}{x}_1 = (x_{11}, x_{12}, \ldots, x_{1n})'$, and

$$\hat{\underset{\sim}{\phi}}_1 = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'(x_{21}, x_{22}, \ldots, x_{2n})'. \tag{2.11}$$

We note that (2.9) is what we would have obtained if we had calculated $p(\underset{\sim}{\phi}|\underset{\sim}{W}_1)$ on the basis of $\underset{\sim}{W}_1$ alone, ignoring $\underset{\sim}{W}_2$.

In the missing observations problem, we are usually concerned with inference on the difference $\delta = \mu_1 - \mu_2$. This can involve use of the predictive distribution

$$h(\underset{\sim}{y}|W_1, W_2) = \int \cdots \int f(\underset{\sim}{y}|\underset{\sim}{\phi}, \underset{\sim}{x}_1^*) p(\underset{\sim}{\phi}|\underset{\sim}{W}_1, \underset{\sim}{W}_2) d\underset{\sim}{\phi},$$

and we see from the above that $p(\underset{\sim}{\phi}|\underset{\sim}{W}_1, \underset{\sim}{W}_2)$ can be replaced by $p(\underset{\sim}{\phi}|\underset{\sim}{W}_1)$ with the same result. (Details are given by Draper and Guttman, 1977).

This seems intuitively reasonable from (2.2a) and (2.2b). The distribution of $\underset{\sim}{y}$ conditional on $\underset{\sim}{x}_1^*$ depends on α, β and d_{11}, and the distribution of $\underset{\sim}{x}_1^*$ depends only on η_1 and d_{22} and provides no information on $\underset{\sim}{\phi} = (\alpha, \beta, d_{11})'$. Thus, in making inferences about $\underset{\sim}{\phi}$. $\underset{\sim}{x}_2^*$ can effectively be ignored.

Example 2.2. In this example, we suppose that $(x_1, \ldots, x_k)$ has the multinomial distribution

$$m_k(x_1, \ldots, x_k) = \frac{\eta!}{x_1! \ldots x_k! (\eta - x_1 - \ldots x_k)!} \mathcal{T}_1^{x_1} \ldots \mathcal{T}_k^{x_k} (1 - \mathcal{T}_1 - \ldots - \mathcal{T}_k)^{\eta - x_1 - \ldots - x_i} \tag{2.12}$$

where $0 \le x_i \le \eta$, $0 \le \sum_1^k x_i \le \eta$, $0 < \mathcal{T}_i < 1$, and $\sum_1^k \mathcal{T}_i = 1$.

Let $(1 \le k_1 \le k-1)$

$$\begin{aligned} \theta_1 &= \Upsilon_1 \\ \theta_2 &= \Upsilon_2 \\ &\vdots \\ \theta_{k_1} &= \Upsilon_{k_1} \\ \phi_1 &= \Upsilon_{k_1+1}/(1-\sum_1^{k_1}\Upsilon_i) \\ &\vdots \\ \phi_{k_2} &= \Upsilon_k/(1-\sum_1^{k_1}\Upsilon_i) \end{aligned} \qquad \text{OR} \qquad \begin{aligned} \Upsilon_1 &= \theta_1 \\ \Upsilon_2 &= \theta_2 \\ &\vdots \\ \Upsilon_{k_1} &= \theta_{k_1} \\ \Upsilon_{k_1+1} &= \phi_1(1-\sum_1^{k_1}\theta_i) \\ &\vdots \\ \Upsilon_k &= \phi_{k_2}(1-\sum_1^{k_1}\theta_i) \end{aligned} \tag{2.13}$$

where $k_2 = k-k_1$.

We note that (2.13), viewed as a tranformation from $(\Upsilon_1,\ldots\Upsilon_k)$ to $(\theta_1,\ldots,\theta_{k_1},\phi_1,\ldots,\phi_{k_2})$, has Jacobian whose absolute value is $(1-\sum_1^{k_1}\theta_i)^{k_2}$.

Now, as is well known, we may write (2.12) as

$$m_k(x_1,\ldots,x_k) = m_k(x_1,\ldots,x_{k_1})f(x_{k_1+1},\ldots,x_k|x_1,\ldots,x_{k_1}) \tag{2.14}$$

where

$$m_{k_1}(x_1,\ldots,x_{k_1}) = \frac{\eta!}{x_1!\ldots x_k!(\eta-x_1-\ldots-x_{k_1})!}\Upsilon_1^{x_1}\ldots\Upsilon_{k_1}^{x_{k_1}}(1-\Upsilon_1-\ldots-\Upsilon_{k_1})^{\eta-x_1-\ldots-x_{k_1}} \tag{2.14a}$$

$$f(x_{k_1+1},\ldots,x_k|x_1,\ldots,x_{k_1}) = \frac{(\eta-x_1-\ldots x_{k_1})!}{x_{k_1+1}!\ldots x_k!(\eta-\sum_1^k x_i)!}\phi_1^{x_{k_1+1}}\ldots\phi_{k_2}^{x_{k_1+k_2}}$$

$$\times(1-\phi_1-\ldots-\phi_{k_2})^{\eta-\sum_1^k x_i} \tag{2.14b}$$

It is often the case that, given $x_1,\ldots,x_{k_1}$, the probability of observing characteristic k_1+j, $j=1,\ldots,k_2$, is of interest and, of course, $\phi_j = \mathcal{T}_{k_1+j}/(1-\sum_1^{k_1} \mathcal{T}_i)$ is this conditional probability. Suppose that in gathering data to estimate $\underset{\sim}{\phi}=(\phi_1,\ldots,\phi_{k_2})'$, that the data on $(x_1,\ldots,x_k)$ has two <u>independent</u> pieces -

$$\underset{\sim}{W}_1 = (x_1^{(1)}, x_2^{(1)},\ldots,x_k^{(1)})',$$

and (2.16)

$$\underset{\sim}{W}_2 = (x_1^{(2)},\ldots,x_{k_1}^{(2)})'$$

where $\underset{\sim}{W}_1$ has probability function given by (2.14) and $\underset{\sim}{W}_2$ has probability function given by (2.14a), with $1 \le k_1 \le k-1$. Suppose too that a-priori,

$$p(\mathcal{T}_1,\ldots,\mathcal{T}_k) \propto c \tag{2.17}$$

Hence, consulting (2.13), and recalling that $\underset{\sim}{\phi}$ is of interest and $\underset{\sim}{\theta}$ is nuisance, we find

$$p(\underset{\sim}{\phi},\underset{\sim}{\theta}) \propto (1-\sum_1^{k_1} \theta_i)^{k_2} \quad . \tag{2.17a}$$

Hence, condition 4 of Theorem 1 is satisfied, and letting $\underset{\sim}{W}_1 = \left(x_1^{(1)},\ldots,x_{k_1}^{(1)} \vdots\; x_{k_1+1}^{(1)},\ldots,x_k^{(1)}\right) = (\underset{\sim}{W}_{11}' \vdots \underset{\sim}{W}_{12}')$, it is easy to see that conditions 1-3 of Theorem 1 also hold. Hence, from Theorem 1, the marginal posterior of $\underset{\sim}{\phi}$, given $\underset{\sim}{W}_1$, $\underset{\sim}{W}_2$, does not depend on $\underset{\sim}{W}_2$. Indeed, it is easy to see directly that the posterior of $\underset{\sim}{\phi}$, given $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$ is

$$p(\underset{\sim}{\phi}|\underset{\sim}{W}_1, \underset{\sim}{W}_2) \propto \phi_1^{x^{(1)}_{k_1+1}} \cdots \phi_{k_2}^{x^{(1)}_{k_1+k_2}} (1-\phi_1-\cdots-\phi_{k_2})^{\eta-\sum_1^k x_i^{(1)}}$$

so that $p(\underset{\sim}{\phi}|\underset{\sim}{W}_1, \underset{\sim}{W}_2) = p(\underset{\sim}{\phi}|\underset{\sim}{W}_1) = p(\underset{\sim}{\phi}|\underset{\sim}{W}_{11}, \underset{\sim}{W}_{12})$.

3. A SUBSIDIARY RESULT.

In this section, we suppose that the data $\underset{\sim}{W} = (\underset{\sim}{W}_1, \underset{\sim}{W}_2)$, with $\underset{\sim}{W}_1$ independent of $\underset{\sim}{W}_2$, is such that

1. The density function of $\underset{\sim}{W}_1$ is such that

$$f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1 | \underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \underset{\sim}{\theta}) = f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1 | \underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2) \tag{3.1}$$

and the density function of W_2, is such that

$$f_{\underset{\sim}{W}_2}(\underset{\sim}{w}_2 | \underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \underset{\sim}{\theta}) = f_{W_2}(\underset{\sim}{w}_2 | \underset{\sim}{\theta}) \tag{3.2}$$

2. $p(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \underset{\sim}{\theta}) \propto a(\underset{\sim}{\theta})\ b(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2)$

We suppose interest is in $\underset{\sim}{\phi}_1$ alone, so that $\underset{\sim}{\theta}$ and $\underset{\sim}{\phi}_2$ are vectors of nuisance parameters. We have

Theorem 2. <u>If conditions 1 and 2 above hold, then the posterior of</u> $\underset{\sim}{\phi}_1$, <u>given</u> $\underset{\sim}{W} = (\underset{\sim}{W}_1, \underset{\sim}{W}_2)$ <u>does not depend on</u> $\underset{\sim}{W}_2$, <u>given that</u> $\underset{\sim}{W}_1$ <u>and</u> $\underset{\sim}{W}_2$ <u>are independent</u>.

<u>Proof</u>. We have, since $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$ are independent, and that conditions 1 and 2 hold, that

$$p(\underset{\sim}{\phi}_1 | \underset{\sim}{W}_1, \underset{\sim}{W}_2) \propto \int_{\underset{\sim}{\phi}_2} \int_{\underset{\sim}{\theta}} a(\underset{\sim}{\theta}) b(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2) f_{\underset{\sim}{W}}(\underset{\sim}{w} | \underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \underset{\sim}{\theta}) d\underset{\sim}{\theta}\ d\underset{\sim}{\phi}_2 \tag{3.3a}$$

$$\propto \int_{\underset{\sim}{\phi}_2} b(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2) f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1 | \underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2) \{\int_{\underset{\sim}{\theta}} a(\underset{\sim}{\theta}) f_{\underset{\sim}{W}_2}(\underset{\sim}{w}_2 | \underset{\sim}{\theta}) d\underset{\sim}{\theta}\}\ d\underset{\sim}{\phi}_2 \tag{3.3b}$$

$$\propto \int_{\underset{\sim}{\phi}_2} b(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2) f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1 | \underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2) d\underset{\sim}{\phi}_2 \tag{3.3c}$$

where the inner integral in (3.3b), a function of $\underset{\sim}{W}_2$ only, has been absorbed in the constant of proportionality, and the theorem is proved.

Note that the posterior of $\underset{\sim}{\phi}_1$, given $\underset{\sim}{W}_1$ alone, need not exist, even though the posterior of $\underset{\sim}{\phi}_1$, given $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$ is independent of $\underset{\sim}{W}_2$. For a direct computation of the posterior of $\underset{\sim}{\phi}_1$, given (only) $\underset{\sim}{W}_1$, yields (under the assumptions of Theorem 2),

$$p(\underset{\sim}{\phi}_1|\underset{\sim}{W}_1) \propto \int_{\underset{\sim}{\phi}_2} \int_{\underset{\sim}{\theta}} a(\underset{\sim}{\theta}) b(\underset{\sim}{\phi}_1,\underset{\sim}{\phi}_2) f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1|\underset{\sim}{\phi}_1,\underset{\sim}{\phi}_2) d\underset{\sim}{\theta}\, d\underset{\sim}{\phi}_2 \tag{3.4a}$$

$$\propto \int_{\underset{\sim}{\phi}} b(\underset{\sim}{\phi}_1,\underset{\sim}{\phi}_2) f_{\underset{\sim}{W}_1}(\underset{\sim}{w}_1|\underset{\sim}{\phi}_1,\underset{\sim}{\phi}_2)\{\int_{\underset{\sim}{\theta}} a(\underset{\sim}{\theta}) d\underset{\sim}{\theta}\}\, d\underset{\sim}{\phi}_2; \tag{3.4b}$$

but if the prior density of $\underset{\sim}{\theta}$, $a(\underset{\sim}{\theta})$ is improper, then $p(\underset{\sim}{\phi}_1|\underset{\sim}{W}_1)$ does not exist. For further remarks on improper priors, see Dawid, Stone, and Zidek (1973).

An example is provided by the following:

$$\underset{\sim}{W}_1 = \{(z_i, x_{2i});\ \ i = 1,\ldots,n)\}, \text{ where the } z_i \text{ are constants,}$$

$$f(x_{2i}|z_i;\alpha,\beta,\tau^2) = (2\pi\tau^2)^{-1/2} \exp - \frac{1}{2\tau^2}(x_{2i}-\alpha-\beta z_i)^2, \tag{3.5}$$

and

$$\underset{\sim}{W}_2 = \{x^*_{1j}, j=1,\ldots,n^*\}$$

$$f(x^*_{1j}|\mu,\sigma^2) = (2\pi\sigma^2)^{-1} \exp - \frac{1}{2\tau^2}(x^*_{1j}-\mu)^2, \tag{3.6}$$

where

$$\underset{\sim}{\phi}_1 = (\alpha,\beta)'; \quad \underset{\sim}{\phi}_2 = \phi_2 = \tau^2; \quad \underset{\sim}{\theta} = (\mu,\sigma^2)' \tag{3.7}$$

Finally, we assume that

$$p(\underset{\sim}{\phi}_1,\phi_2,\underset{\sim}{\theta}) = a(\underset{\sim}{\theta})b(\underset{\sim}{\phi}_1,\phi_2), \tag{3.8}$$

with

$$a(\underset{\sim}{\theta}) = a(\mu,\sigma^2) \propto 1/\sigma^2 \tag{3.8a}$$

and

$$b(\underset{\sim}{\phi}_1,\phi_2) = b(d,\beta,\tau^2)$$

$$= \frac{|\underset{\sim}{A}|^{1/2} S_0^{(n_0-2)/2}}{2^{n_0/2}(\Gamma(\tfrac{1}{2}))^2\Gamma(\frac{n_0-2}{2})} (\tau^2)^{-(\frac{1}{2}n_0+1)} \quad \exp\{-\frac{1}{2\tau^2}[S_0+Q]\} \tag{3.8b}$$

where

$$Q = (\alpha-\alpha_0,\beta-\beta_0)\underset{\sim}{A}(\alpha-\alpha_0,\beta-\beta_0)' \tag{3.8c}$$

with $\underset{\sim}{A}$ positive definite. It turns out that the posterior of (α,β), given W_1 and W_2, is connected to the distribution of a bivariate t, degrees of

freedom ($n+n_0-2$), which is independent of the x^*_{1j}'s, that is, of $\underset{\sim}{W}_2$, and, further, that the posterior of $\underset{\sim}{\phi}_1$, given $\underset{\sim}{W}_1$ alone, does not exist - these results may be seen by substituting in (3.3c) and (3.4b).

ACKNOWLEDGEMENTS

This work was partially supported by the United States Army under Contracts DAAF29-80-C-0041 and DAAG29-80-C-0113, by NSERC (Canada) under Grant No. A8743, and by the Canadian Council in Humanities and Social Sciences Under Award No. W760082. We are grateful to Mervyn Stone for valuable discussion.

REFERENCES

Dawid, A.P., Stone, M. and Zidek, J.V. (1973). Marginalization paradoxes in Bayesian and structural inference. Journal of the Royal Statistical Society, B-35, 189-213 (discussion 214-233).

Draper, N.R. and Guttman, I. (1977). Inference from incomplete data on the difference of means of correlated variables: a Bayesian approach. University of Wisconsin Statistics Department Technical Report No. 496, August.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER #2409	2. GOVT ACCESSION NO. AD-A120 958	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Dropping Observations Without Affecting Posterior and Predictive Distributions		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Norman R. Draper and Irwin Guttman		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041 DAAG29-80-C-0113
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE August 1982 13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bayes' theorem, Nuisance parameters, Predictive distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Suppose in a distribution problem, the sample information $\underset{\sim}{W}$ is split into two pieces $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$, and the parameters involved are split into two sets, $\underset{\sim}{\phi}$ containing the parameters of interest, and $\underset{\sim}{\theta}$ containing nuisance parameters. It is shown that, under certain conditions, the posterior

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (continued)

W sub 2 prime

distribution of $\underset{\sim}{\phi}$ does not depend on the data $\underset{\sim}{W}_2$, which can thus be ignored. This also has consequences for the predictive distribution of future (or missing) observations. In fact, under similar conditions, the predictive distributions using $\underset{\sim}{W}$ or just $\underset{\sim}{W}_1$ are identical.

W sub 1